for SEQ ID NO:16

RESULT 1
AA142913
LOCUS AA142913 572 bp mRNA linear EST 14-MAY-1997
DEFINITION z140e07.s1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone IMAGE:504420 3', mRNA sequence.
ACCESSION AA142913
VERSION AA142913.1 GI:1712300
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 572)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

Sequence Mail Library
for 09/873367



IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1212 Std Error: 0.00
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 442.

FEATURES Location/Qualifiers
source 1..572
/organism="Homo sapiens"
/db_xref="GDB:3809408"
/db_xref="taxon:9606"
/clone="IMAGE:504420"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGCGGCCGCCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by M. Fatima Bonaldo."

BASE COUNT 191 a 87 c 96 g 198 t
ORIGIN

Query Match 100.0%; Score 572; DB 9; Length 572;
Best Local Similarity 100.0%; Pred. No. 6.5e-94;
Matches 572; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TTTTTTGACATTGTTCTACTGTTTTATTGACTCGTTGCATTTACAAGTTTTGCTAATGAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTGACATTGTTCTACTGTTTTATTGACTCGTTGCATTTACAAGTTTTGCTAATGAT 60

Qy     61 ACACAGTCTACACTTACTAATAAATTATACTCACAGTGTTTTTAGTGATGTGACTTTGTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ACACAGTCTACACTTACTAATAAATTATACTCACAGTGTTTTTAGTGATGTGACTTTGTT 120

Qy    121 TCAATATTTTATAATAAAAGATTATAGGAGTAATTACAGACAATGATAGAAAAGTTTGAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TCAATATTTTATAATAAAAGATTATAGGAGTAATTACAGACAATGATAGAAAAGTTTGAG 180

Qy    181 GCATCGTGACAAAATAGTGCAAAAGCCTAAGTTATCCAAAAGATGTAGTGATCATAATTA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCATCGTGACAAAATAGTGCAAAAGCCTAAGTTATCCAAAAGATGTAGTGATCATAATTA 240

Qy    241 TAAAGACTGTGTAGTGTCCCTGGGAAATGCTTACAATGAGATACCAAGCAGTCAAAACGG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TAAAGACTGTGTAGTGTCCCTGGGAAATGCTTACAATGAGATACCAAGCAGTCAAAACGG 300

Qy    301 AATCTAACCACGCACCTGTACAGTAGTTACAAAGGTATTACAAAGCTTGTCTCTGCATGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AATCTAACCACGCACCTGTACAGTAGTTACAAAGGTATTACAAAGCTTGTCTCTGCATGA 360

Qy    361 ACACAGTAAAGAAGTCACACATACACAAACGACTACAATGGTGTTCTGGTATTGCGACTG 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ACACAGTAAAGAAGTCACACATACACAAACGACTACAATGGTGTTCTGGTATTGCGACTG 420

Qy    421 TTTGTTTTTTCTTCTTTAAATATTATTTTGCTTTATTGTTGTAATGTTATTTTTGTAATA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TTTGTTTTTTCTTCTTTAAATATTATTTTGCTTTATTGTTGTAATGTTATTTTTGTAATA 480

Qy    481 AATAAATTCAGAGAGAACATCCTACTATTAGACAAGGAAAATGCCAGAAATCTGAGATAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AATAAATTCAGAGAGAACATCCTACTATTAGACAAGGAAAATGCCAGAAATCTGAGATAT 540

Qy    541 TTTCCCTCTTATGGCCGTATTATATTGGTTAC 572
          ||||||||||||||||||||||||||||||
Db    541 TTTCCCTCTTATGGCCGTATTATATTGGTTAC 572
```

for SEQ ID NO:87

RESULT 2
HUMRSC338
LOCUS HUMRSC338 2416 bp mRNA linear PRI 06-OCT-2001
DEFINITION Human mRNA for KIAA0001 gene, complete cds.
ACCESSION D13626
VERSION D13626.1 GI:285994
KEYWORDS KIAA0001.
SOURCE Homo sapiens male myeloblast cell_line KG-1 cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gerard,C. and Gerard,N.P.
TITLE C5A anaphylatoxin and its seven transmembrane-segment receptor
JOURNAL Annu. Rev. Immunol. 12, 775-808 (1994)
MEDLINE 94280786
REFERENCE 2
AUTHORS Nomura,N., Miyajima,N., Sazuka,T., Tanaka,A., Kawarabayasi,Y., Sato,S., Nagase,T., Seki,N., Ishikawa,K. and Tabata,S.
TITLE Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1
JOURNAL DNA Res. 1 (1), 27-35 (1994)
MEDLINE 96051387
REFERENCE 3
AUTHORS Nomura,N., Miyajima,N., Sazuka,T., Tanaka,A., Kawarabayasi,Y., Sato,S., Nagase,T., Seki,N., Ishikawa,K. and Tabata,S.
TITLE Prediction of the coding sequences of unidentified human genes. I. The coding sequences of 40 new genes (KIAA0001-KIAA0040) deduced by analysis of randomly sampled cDNA clones from human immature myeloid cell line KG-1 (supplement)
JOURNAL DNA Res. 1 (1), 47-56 (1994)
MEDLINE 96051389
REFERENCE 4 (bases 1 to 2416)
AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Nomura,N.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1992) Osamu Ohara, Kazusa DNA Research Institute; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913)

FEATURES Location/Qualifiers
source 1. .2416
/organism="Homo sapiens"
/db_xref="taxon:9606"
/sex="male"
/cell_line="KG-1"
/cell_type="myeloblast"
gene 1. .2416
/gene="KIAA0001"
5'UTR <1. .216
/gene="KIAA0001"
CDS 217. .1233
/gene="KIAA0001"
/codon_start=1
/protein_id="BAA02791.1"
/db_xref="GI:285995"
/translation="MINSTSTQPPDESCSQNLLITQQIIPVLYCMVFIAGILLNG
WIFFYVPSSKSFIIYLKNIVIADFVMSLTFPFKILGDSGLGPWQLNVFVCRVSAV
VNMYVSIVFFGLISFDRYYKIVKPLWTSFIQSVSYSKLLSVIVWMLMLLAVPNI
NQSVREVTQIKCIELKSELGRKWHKASNYIFVAIFWIVFLLLIVFYTAITKKIFK
KSSRNSTSVKKKSSRNIFSIVFVFFVCFVPYHIARIPYTKSQTEAHYSCQSKEIL
KEFTLLLSAANVCLDPIIYFFLCQPFREILCKKLHIPLKAQNDLDISRIKRGNTTI
TDTL"
3'UTR 1234. .>2416
/gene="KIAA0001"

BASE COUNT 782 a 476 c 405 g 753 t
ORIGIN

Query Match 100.0%; Score 2416; DB 9; Length 2416;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2416; Conservative 0; Mismatches 0; Indels 0; Gaps

```
Qy       1 GAACAGTGTTACCTTGGAGCCTACAATGAGAGGTATTTCAAAATGAGTGAAGCATGACTC 60
Db       1 GAACAGTGTTACCTTGGAGCCTACAATGAGAGGTATTTCAAAATGAGTGAAGCATGACTC 60

Qy      61 TCACAGATGAAGGCCTAGACGCAGGATCTTTAATGGAAAAACACTTGGGCCACTTCAAGA 120
Db      61 TCACAGATGAAGGCCTAGACGCAGGATCTTTAATGGAAAAACACTTGGGCCACTTCAAGA 120

Qy     121 CGACAAACGCTCACTGGGCAAAACACCTTCACTGAAAAGAGACCTCATATTATGCAAAAA 180
Db     121 CGACAAACGCTCACTGGGCAAAACACCTTCACTGAAAAGAGACCTCATATTATGCAAAAA 180

Qy     181 AAATCTTAAGAGGCCTCTGCCTTCAGAAGTTACAAGATGATCAATTCAACCTCCACACAG 240
Db     181 AAATCTTAAGAGGCCTCTGCCTTCAGAAGTTACAAGATGATCAATTCAACCTCCACACAG 240

Qy     241 CCTCCAGATGAATCCTGCTCTCAGAACCTCCTGATCACTCAGCAGATCATTCCTGTGCTG 300
Db     241 CCTCCAGATGAATCCTGCTCTCAGAACCTCCTGATCACTCAGCAGATCATTCCTGTGCTG 300

Qy     301 TACTGTATGGTCTTCATTGCGGGAATCCTACTCAATGGAGTGTCAGGATGGATATTCTTT 360
Db     301 TACTGTATGGTCTTCATTGCGGGAATCCTACTCAATGGAGTGTCAGGATGGATATTCTTT 360

Qy     361 TACGTGCCCAGCTCTAAGAGTTTCATCATCTATCTCAAGAACATTGTTATTGCTGACTTT 420
Db     361 TACGTGCCCAGCTCTAAGAGTTTCATCATCTATCTCAAGAACATTGTTATTGCTGACTTT 420

Qy     421 GTGATGAGCCTGACTTTTCCTTTCAAGATCCTTGGTGACTCAGGCCTTGGTCCCTGGCAG 480
Db     421 GTGATGAGCCTGACTTTTCCTTTCAAGATCCTTGGTGACTCAGGCCTTGGTCCCTGGCAG 480

Qy     481 CTGAACGTGTTTGTGTGCAGGGTCTCTGCCGTGCTCTTCTACGTCAACATGTACGTCAGC 540
Db     481 CTGAACGTGTTTGTGTGCAGGGTCTCTGCCGTGCTCTTCTACGTCAACATGTACGTCAGC 540

Qy     541 ATTGTGTTCTTTGGGCTCATCAGCTTTGACAGGTATTATAAAATTGTAAAGCCTCTTTGG 600
Db     541 ATTGTGTTCTTTGGGCTCATCAGCTTTGACAGGTATTATAAAATTGTAAAGCCTCTTTGG 600

Qy     601 ACTTCTTTCATCCAGTCAGTGAGTTACAGCAAACTTCTGTCAGTGATAGTATGGATGCTC 660
```

```
Db    601 ACTTCTTTCATCCAGTCAGTGAGTTACAGCAAACTTCTGTCAGTGATAGTATGGATGCTC 660
Qy    661 ATGCTCCTCCTTGCTGTTCCAAATATTATTCTCACCAACCAGAGTGTTAGGGAGGTTACA 720
Db    661 ATGCTCCTCCTTGCTGTTCCAAATATTATTCTCACCAACCAGAGTGTTAGGGAGGTTACA 720
Qy    721 CAAATAAAATGTATAGAACTGAAAAGTGAACTGGGACGGAAGTGGCACAAAGCATCAAAC 780
Db    721 CAAATAAAATGTATAGAACTGAAAAGTGAACTGGGACGGAAGTGGCACAAAGCATCAAAC 780
Qy    781 TACATCTTCGTGGCCATCTTCTGGATTGTGTTTCTTTTGTTAATCGTTTTCTATACTGCT 840
Db    781 TACATCTTCGTGGCCATCTTCTGGATTGTGTTTCTTTTGTTAATCGTTTTCTATACTGCT 840
Qy    841 ATCACAAAGAAAATCTTTAAGTCCCACCTTAAGTCAAGTCGGAATTCCACTTCGGTCAAA 900
Db    841 ATCACAAAGAAAATCTTTAAGTCCCACCTTAAGTCAAGTCGGAATTCCACTTCGGTCAAA 900
Qy    901 AAGAAATCTAGCCGCAACATATTCAGCATCGTGTTTGTGTTTTTTGTCTGTTTTGTACCT 960
Db    901 AAGAAATCTAGCCGCAACATATTCAGCATCGTGTTTGTGTTTTTTGTCTGTTTTGTACCT 960
Qy    961 TACCATATTGCCAGAATCCCCTACACAAAGAGTCAGACCGAAGCTCATTACAGCTGCCAG 1020
Db    961 TACCATATTGCCAGAATCCCCTACACAAAGAGTCAGACCGAAGCTCATTACAGCTGCCAG 1020
Qy   1021 TCAAAAGAAATCTTGCGGTATATGAAAGAATTCACTCTGCTACTATCTGCTGCAAATGTA 1080
Db   1021 TCAAAAGAAATCTTGCGGTATATGAAAGAATTCACTCTGCTACTATCTGCTGCAAATGTA 1080
Qy   1081 TGCTTGGACCCTATTATTTATTTCTTTCTATGCCAGCCGTTTAGGGAAATCTTATGTAAG 1140
Db   1081 TGCTTGGACCCTATTATTTATTTCTTTCTATGCCAGCCGTTTAGGGAAATCTTATGTAAG 1140
Qy   1141 AAATTGCACATTCCATTAAAAGCTCAGAATGACCTAGACATTTCCAGAATCAAAAGAGGA 1200
Db   1141 AAATTGCACATTCCATTAAAAGCTCAGAATGACCTAGACATTTCCAGAATCAAAAGAGGA 1200
Qy   1201 AATACAACACTTGAAAGCACAGATACTTTGTGAGTTCCTACCCTCTTCCAAAGAAAGACC 1260
Db   1201 AATACAACACTTGAAAGCACAGATACTTTGTGAGTTCCTACCCTCTTCCAAAGAAAGACC 1260
Qy   1261 ACGTGTGCATGTTGTCATCTTCAATTACATAACAGAAATCAATAAGATATGTGCCCTCAT 1320
Db   1261 ACGTGTGCATGTTGTCATCTTCAATTACATAACAGAAATCAATAAGATATGTGCCCTCAT 1320
Qy   1321 CATAAATATCATCTCTAGCACTGCCATCCAATTTAGTTCAATAAAATTCAAATATAAGTT 1380
Db   1321 CATAAATATCATCTCTAGCACTGCCATCCAATTTAGTTCAATAAAATTCAAATATAAGTT 1380
Qy   1381 TCCATGCTTTTTTGTAACATCAAAGAAAACATACCCATCAGTAATTTCTCTAATACTGAC 1440
Db   1381 TCCATGCTTTTTTGTAACATCAAAGAAAACATACCCATCAGTAATTTCTCTAATACTGAC 1440
Qy   1441 CTTTCTATTCTCTATTAATAAAAAATTAATACATACAATTATTCAATTCTATTATATTAA 1500
Db   1441 CTTTCTATTCTCTATTAATAAAAAATTAATACATACAATTATTCAATTCTATTATATTAA 1500
Qy   1501 AATAAGTTAAAGTTTATAACCACTAGTCTGGTCAGTTAATGTAGAAATTTAAATAGTAAA 1560
Db   1501 AATAAGTTAAAGTTTATAACCACTAGTCTGGTCAGTTAATGTAGAAATTTAAATAGTAAA 1560
Qy   1561 TAAAACACAACATAATCAAAGACAACTCACTCAGGCATCTTCTTTCTCTAAATACCAGAA 1620
Db   1561 TAAAACACAACATAATCAAAGACAACTCACTCAGGCATCTTCTTTCTCTAAATACCAGAA 1620
Qy   1621 TCTAGTATGTAATTGTTTTCAACACTGTCCTTAAAGACTAACTTGAAAGCAGGCACAGTT 1680
Db   1621 TCTAGTATGTAATTGTTTTCAACACTGTCCTTAAAGACTAACTTGAAAGCAGGCACAGTT 1680
Qy   1681 TGATGAAGGGCTAGAGAGCTGTTTGCAATAAAAAGTCAGGTTTTTTTCCTGATTTGAAGA 1740
Db   1681 TGATGAAGGGCTAGAGAGCTGTTTGCAATAAAAAGTCAGGTTTTTTTCCTGATTTGAAGA 1740
Qy   1741 AGCAGGAAAAGCTGACACCCAGACAATCACTTAAGAAACCCCTTATTGATGTATTTCATG 1800
Db   1741 AGCAGGAAAAGCTGACACCCAGACAATCACTTAAGAAACCCCTTATTGATGTATTTCATG 1800
Qy   1801 GCACTGCAAAGGAAGAGGAATATTAATTGTATACTTAGCAAGAAAATTTTTTTTTTCTGA 1860
Db   1801 GCACTGCAAAGGAAGAGGAATATTAATTGTATACTTAGCAAGAAAATTTTTTTTTTCTGA 1860
Qy   1861 TAGCACTTTGAGGATATTAGATACATGCTAAATATGTTTTCTACAAAGACTTACGTCATT 1920
Db   1861 TAGCACTTTGAGGATATTAGATACATGCTAAATATGTTTTCTACAAAGACTTACGTCATT 1920
Qy   1921 TAATGAGCCTGGGGTTCTGGTGTTAGAATATTTTTAAGTAGGCTTTACTGAGAGAAACTA 1980
Db   1921 TAATGAGCCTGGGGTTCTGGTGTTAGAATATTTTTAAGTAGGCTTTACTGAGAGAAACTA 1980
Qy   1981 AATATTGGCATACGTTATCAGCAACTTCCCCTGTTCAATAGTATGGGAAAAATAAGATGA 2040
Db   1981 AATATTGGCATACGTTATCAGCAACTTCCCCTGTTCAATAGTATGGGAAAAATAAGATGA 2040
Qy   2041 CTGGGAAAAAGACACACCCACACCGTAGAACATATATTAATCTACTGGCGAATGGGAAAG 2100
Db   2041 CTGGGAAAAAGACACACCCACACCGTAGAACATATATTAATCTACTGGCGAATGGGAAAG 2100
Qy   2101 GAGACCATTTTCTTAGAAAGCAAATAAACTTGATTTTTTTAAATCTAAAATTTACATTAA 2160
Db   2101 GAGACCATTTTCTTAGAAAGCAAATAAACTTGATTTTTTTAAATCTAAAATTTACATTAA 2160
Qy   2161 TGAGTGCAAAATAACACATAAAATGAAAATTCACACATCACATTTTTCTGGAAAACAGAC 2220
Db   2161 TGAGTGCAAAATAACACATAAAATGAAAATTCACACATCACATTTTTCTGGAAAACAGAC 2220
Qy   2221 GGATTTTACTTCTGGAGACATGGCATACGGTTACTGACTTATGAGCTACCAAAACTAAAT 2280
Db   2221 GGATTTTACTTCTGGAGACATGGCATACGGTTACTGACTTATGAGCTACCAAAACTAAAT 2280
Qy   2281 TCTTTCTCTGCTATTAACTGGCTAGAAGACATTCATCTATTTTTCAAATGTTCTTTCAAA 2340
Db   2281 TCTTTCTCTGCTATTAACTGGCTAGAAGACATTCATCTATTTTTCAAATGTTCTTTCAAA 2340
Qy   2341 ACATTTTTATAAGTAATGTTTGTATCTATTTCATGCTTTACTGTCTATATACTAATAAAG 2400
Db   2341 ACATTTTTATAAGTAATGTTTGTATCTATTTCATGCTTTACTGTCTATATACTAATAAAG 2400
Qy   2401 AAATGTTTTAATACTG 2416
Db   2401 AAATGTTTTAATACTG 2416
```

for SEQ ID NO: 453

RESULT 3
AW006758
LOCUS AW006758 362 bp mRNA linear EST 08-MAR-2000
DEFINITION wt07a10.x1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:2506746 3', mRNA sequence.
ACCESSION AW006758
VERSION AW006758.1 GI:5855536
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 362)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 442 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES Location/Qualifiers
source 1. .362
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2506746"
/clone_lib="NCI_CGAP_Co3"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization."

BASE COUNT 82 a 93 c 91 g 96 t
ORIGIN

Query Match 100.0%; Score 330; DB 10; Length 362;
Best Local Similarity 100.0%; Pred. No. 6.3e-89;
Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TGTGGATAAATATATTAGCAAATAAATATATTTCTTAACATAGTGCCTGATTCAAGCGTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TGTGGATAAATATATTAGCAAATAAATATATTTCTTAACATAGTGCCTGATTCAAGCGTC 60

Qy     61 TGTCTGGTTCAGATATAAATACCCATGTGGGTACCTAGGTGCTAGTCTCCCCACTAACTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TGTCTGGTTCAGATATAAATACCCATGTGGGTACCTAGGTGCTAGTCTCCCCACTAACTG 120

Qy    121 AGGGAAAAAGGTTCCCAGGTGGGGTCCTCTGCCCACTTTGCCACCACATTCACATTCCAA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGGGAAAAAGGTTCCCAGGTGGGGTCCTCTGCCCACTTTGCCACCACATTCACATTCCAA 180

Qy    181 ATGGGATAATGCCTGAGGGGCCAAGAGTGGTCAGGCTGCCCTGGGGTGAATGTCACCCTG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ATGGGATAATGCCTGAGGGGCCAAGAGTGGTCAGGCTGCCCTGGGGTGAATGTCACCCTG 240

Qy    241 ATGAGGCCCATCAGCTCTTGTCCACTCAGTGAGGCCAGACTTGTGCTCTAATCCACTCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ATGAGGCCCATCAGCTCTTGTCCACTCAGTGAGGCCAGACTTGTGCTCTAATCCACTCTC 300

Qy    301 CTGTGGGTCCCTGGCCTGTATGGCTTATAC 330
          |||||||||||||||||||||||||||||
Db    301 CTGTGGGTCCCTGGCCTGTATGGCTTATAC 330
```

for SEQ ID NO: 462

RESULT 4
ABV29346
ID ABV29346 standard; cDNA; 734 BP.
XX
AC ABV29346;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 29337.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer
XX
PS Claim 1; Page 6285; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ Sequence 734 BP; 193 A; 157 C; 114 G; 218 T; 52 other;

Query Match 100.0%; Score 320; DB 23; Length 734;
Best Local Similarity 100.0%; Pred. No. 1.1e-82;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CACAGTGGCAGATTTTCTTTAATAGATATATTTCAAACAGATACAACAAATTAAAAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 CACAGTGGCAGATTTTCTTTAATAGATATATTTCAAACAGATACAACAAATTAAAAAATC 152

Qy     61 TAATTCACGGCAGGTAAACATGGGTGCTCAAAAAGTTCCACATAGACATTTACACTTGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 TAATTCACGGCAGGTAAACATGGGTGCTCAAAAAGTTCCACATAGACATTTACACTTGGG 212

Qy    121 CCATCAGTATTTCCCTCACATTCCCTTTTGTTAAGTCCCATCTTCGCAGTGGCAGTACAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    213 CCATCAGTATTTCCCTCACATTCCCTTTTGTTAAGTCCCATCTTCGCAGTGGCAGTACAG 272

Qy    181 GAGAAATCTCCACCGTCACCGCACAATCCACCAGGCGCATTACCACCTGAAGTGAAGGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 GAGAAATCTCCACCGTCACCGCACAATCCACCAGGCGCATTACCACCTGAAGTGAAGGTC 332

Qy    241 TCATCTCGAAGGTGCGCTCAGCCATAAAAAGAAAACATATTACAGAAAGGAAAAATAAGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 TCATCTCGAAGGTGCGCTCAGCCATAAAAAGAAAACATATTACAGAAAGGAAAAATAAGT 392

Qy    301 GTGCTCTTCCCACCCGCAAA 320
          ||||||||||||||||||||
Db    393 GTGCTCTTCCCACCCGCAAA 412
```

for SEQ ID NO:462

RESULT 4
AI368176
LOCUS AI368176 392 bp mRNA linear EST 13-FEB-1999
DEFINITION qk65b10.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1873819 3', mRNA sequence.
ACCESSION AI368176
VERSION AI368176.1 GI:4137921
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 392)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1195 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 381.

FEATURES Location/Qualifiers
source 1. .392
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1873819"
/clone_lib="NCI_CGAP_Co8"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 121 a 102 c 65 g 104 t
ORIGIN

Query Match 100.0%; Score 320; DB 9; Length 392;
Best Local Similarity 100.0%; Pred. No. 8.6e-66;
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CACAGTGGCAGATTTTCTTTAATAGATATATTTCAAACAGATACAACAAATTAAAAAATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      6 CACAGTGGCAGATTTTCTTTAATAGATATATTTCAAACAGATACAACAAATTAAAAAATC 65

Qy     61 TAATTCACGGCAGGTAAACATGGGTGCTCAAAAAGTTCCACATAGACATTTACACTTGGG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     66 TAATTCACGGCAGGTAAACATGGGTGCTCAAAAAGTTCCACATAGACATTTACACTTGGG 125

Qy    121 CCATCAGTATTTCCCTCACATTCCCTTTTGTTAAGTCCCATCTTCGCAGTGGCAGTACAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    126 CCATCAGTATTTCCCTCACATTCCCTTTTGTTAAGTCCCATCTTCGCAGTGGCAGTACAG 185

Qy    181 GAGAAATCTCCACCGTCACCGCACAATCCACCAGGCGCATTACCACCTGAAGTGAAGGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 GAGAAATCTCCACCGTCACCGCACAATCCACCAGGCGCATTACCACCTGAAGTGAAGGTC 245

Qy    241 TCATCTCGAAGGTGCGCTCAGCCATAAAAAGAAAACATATTACAGAAAGGAAAAATAAGT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 TCATCTCGAAGGTGCGCTCAGCCATAAAAAGAAAACATATTACAGAAAGGAAAAATAAGT 305

Qy    301 GTGCTCTTCCCACCCGCAAA 320
          ||||||||||||||||||||
Db    306 GTGCTCTTCCCACCCGCAAA 325
```

for SEQ ID NO: 468

RESULT 2
AA629913
LOCUS AA629913 467 bp mRNA linear EST 06-MAR-1998
DEFINITION ad45e02.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone IMAGE:884666 3' similar to contains Alu repetitive element;contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION AA629913
VERSION AA629913.1 GI:2552524
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 467)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 742 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES Location/Qualifiers
source 1. .467
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:884666"
/clone_lib="Stratagene lung carcinoma 937218"
/tissue_type="lung carcinoma"
/cell_line="NCI-H69"
/dev_stage="cell line NCI-H69"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Small cell carcinoma cell line NCI-H69. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

BASE COUNT 98 a 110 c 114 g 145 t
ORIGIN

Query Match 96.5%; Score 453.6; DB 9; Length 467;
Best Local Similarity 98.9%; Pred. No. 5.3e-86;
Matches 456; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      10 GGAGGTTTTATTATTTTTATTTATTTTTGAGACTGAGTCTTGCTCTGTCATCAGGCTGGA 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7 GGAGGTTTTATTATTTTTATTTATTTTTGAGACTGAGTCTTGCTCTGTCATCAGGCTGGA 66

Qy      70 GTGCAGTGGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCAGCCT 129
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      67 GTGCAGTGGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCAGCCT 126

Qy     130 CCCTAGTAGCTGGGATTACAGGTGTCCACCACCATGCCCAATTAATTTTTGTATTTTTGG 189
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     127 CCCTAGTAGCTGGGATTACAGGTGTCCACCACCATGCCCAATTAATTTTTGTATTTTTGG 186

Qy     190 TACAGACAGGGTTTCACCATGTTGGCCAGGATGGTCTCGATCTCGTTGACCTTGTGATCC 249
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     187 TACAGACAGGGTTTCACCATGTTGGCCAGGATGGTCTCGATCTCGTTGACCTTGTGATCC 246

Qy     250 GCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAACCNGCCCCTGGACTACT 309
           |||||||||||||||||||||||||||||||||||     | ||||||||||||||
Db     247 GCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGACGACCGCCCCTGGACTACT 306

Qy     310 TATGGAGGTTTTAAAAAATCTTTTAAGTCCAGGCCTGACGTTTAGAGAAGGTTACAAAGG 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     307 TATGGAGGTTTTAAAAAATCTTTTAAGTCCAGGCCTGACGTTTAGAGAAGGTTACAAAGG 366

Qy     370 CGGCCAGGATCTGAGTATTTCCAAAAAGCTCTGGAGGCAGCATTGAGGTTTCCTTCCAGT 429
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     367 CGGCCAGGATCTGAGTATTTCCAAAAAGCTCTGGAGGCAGCATTGAGGTTTCCTTCCAGT 426

Qy     430 TGAATCACTGACTTTAGGTCGACTGGGGTACTTTGGGTTTT 470
           |||||||||||||||||||||||||||||||||||||||||
Db     427 TGAATCACTGACTTTAGGTCGACTGGGGTACTTTGGGTTTT 467
```

```
RESULT 4                SEQ ID NO:651                 EARLIEST GenBank date 1989
HUMXYPFLA
LOCUS       HUMXYPFLA                1971 bp    mRNA    linear   PRI 29-MAY-2002
DEFINITION  Homo sapiens mRNA for cytochrome P-450 HFLa, complete cds.
ACCESSION   D00408
VERSION     D00408.1  GI:220148
KEYWORDS    CYP3A6; cytochrome P-450; human fetal liver cytochrome P-450.
SOURCE      Homo sapiens fetus liver cDNA to mRNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1971)
  AUTHORS   Komori,M., Nishio,K., Ohi,H., Kitada,M. and Kamataki,T.
  TITLE     Molecular cloning and sequence analysis of cDNA containing the
            entire coding region for human fetal liver cytochrome P-450
  JOURNAL   J. Biochem. 105 (2), 161-163 (1989)
  MEDLINE   89255154
COMMENT     Submitted in computer readable form by M. Komori on 15-Apr-1989.
            The deduced N-terminal amino acid sequence was identical to that of
            P-450 HFLa.
FEATURES             Location/Qualifiers
     source          1. .1971
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="liver"
                     /dev_stage="fetus"
                     /note="106 bp upstream of HindIII site.-clones
                     lambda-HFL33, lambda -HFL10"
     CDS             4. .1515
                     /codon_start=1
                     /product="cytochrome P-450 HFLa"
                     /protein_id="BAA00310.1"
                     /db_xref="GI:220149"
                     /translation="MDLIPNLAVETWLLLAVSLILLYLYGTRTHGLFKKLGIPGPTPL
                     PFLGNALSFRKGYWTFDMECYKKYRKVWGIYDCQQPMLAITDPDMIKTVLVKECYSVF
                     TNRRPFGPVGFMKNAISIAEDEEWKRIRSLLSPTFTSGKLKEMVPIIAQYGDVLVRNL
                     RREAETGKPVTLKHVFGAYSMDVITSTSFGVSIDSLNNPQDPFVENTKKLLRFNPLDP
                     FVLSIKVFPFLTPILEALNITVFPRKVISFLTKSVKQIKEGRLKETQKHRVDFLQLMI
                     DSQNSKDSETHKALSDLELMAQSIIFIFAGYETTSSVLSFIIYELATHPDVQQKVQKE
                     IDTVLPNKAPPTYDTVLQLEYLDMVVNETLRLFPVAMRLERVCKKDVEINGMFIPKGV
                     VVMIPSYVLHHDPKYWTEPEKFLPERFSKKNKDNIDPYIYTPFGSGPRNCIGMRFALV
```

```
                    NMKLALVRVLQNFSFKPCKETQIPLKLRFGGLLLTEKPIVLKAESRDETVSGA"
BASE COUNT      597 a    428 c    397 g    549 t
ORIGIN

  Query Match           100.0%;  Score 1971;  DB 9;  Length 1971;
  Best Local Similarity 100.0%;  Pred. No. 0;
  Matches 1971;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GTGATGGATCTCATCCCAAACTTGGCCGTGGAAACCTGGCTTCTCCTGGCTGTCAGCCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GTGATGGATCTCATCCCAAACTTGGCCGTGGAAACCTGGCTTCTCCTGGCTGTCAGCCTG 60

Qy     61 ATACTCCTCTATCTATATGGAACCCGTACACATGGACTTTTTAAGAAGCTTGGAATTCCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ATACTCCTCTATCTATATGGAACCCGTACACATGGACTTTTTAAGAAGCTTGGAATTCCA 120

Qy    121 GGGCCCACACCTCTGCCTTTTTTGGGAAATGCTTTGTCCTTCCGTAAGGGCTATTGGACG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGGCCCACACCTCTGCCTTTTTTGGGAAATGCTTTGTCCTTCCGTAAGGGCTATTGGACG 180

Qy    181 TTTGACATGGAATGTTATAAAAAGTATAGAAAAGTCTGGGGTATTTATGACTGTCAACAG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTTGACATGGAATGTTATAAAAAGTATAGAAAAGTCTGGGGTATTTATGACTGTCAACAG 240

Qy    241 CCTATGCTGGCTATCACAGATCCCGACATGATCAAAACAGTGCTAGTGAAAGAATGTTAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTATGCTGGCTATCACAGATCCCGACATGATCAAAACAGTGCTAGTGAAAGAATGTTAT 300

Qy    301 TCTGTCTTCACAAACCGGAGGCCTTTCGGGCCAGTGGGATTTATGAAAAATGCCATCTCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCTGTCTTCACAAACCGGAGGCCTTTCGGGCCAGTGGGATTTATGAAAAATGCCATCTCT 360

Qy    361 ATAGCTGAGGATGAAGAATGGAAGAGAATACGATCATTGCTGTCTCCAACATTCACCAGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATAGCTGAGGATGAAGAATGGAAGAGAATACGATCATTGCTGTCTCCAACATTCACCAGC 420

Qy    421 GGAAAACTCAAGGAGATGGTCCCTATCATTGCCCAGTATGGAGATGTGTTGGTGAGAAAT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGAAAACTCAAGGAGATGGTCCCTATCATTGCCCAGTATGGAGATGTGTTGGTGAGAAAT 480

Qy    481 CTGAGGCGGGAAGCAGAGACAGGCAAGCCTGTCACCTTGAAACACGTCTTTGGGGCCTAC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTGAGGCGGGAAGCAGAGACAGGCAAGCCTGTCACCTTGAAACACGTCTTTGGGGCCTAC 540

Qy    541 AGCATGGATGTGATCACTAGCACATCATTTGGAGTGAGCATCGACTCTCTCAACAATCCA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGCATGGATGTGATCACTAGCACATCATTTGGAGTGAGCATCGACTCTCTCAACAATCCA 600

Qy    601 CAAGACCCCTTTGTGGAAAACACCAAGAAGCTTTTAAGATTTAATCCATTAGATCCATTC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CAAGACCCCTTTGTGGAAAACACCAAGAAGCTTTTAAGATTTAATCCATTAGATCCATTC 660

Qy    661 GTTCTCTCAATAAAAGTCTTTCCATTCCTTACCCCAATTCTTGAAGCATTAAATATCACT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTTCTCTCAATAAAAGTCTTTCCATTCCTTACCCCAATTCTTGAAGCATTAAATATCACT 720

Qy    721 GTGTTTCCAAGAAAAGTTATAAGTTTTCTAACAAAATCTGTAAAACAGATAAAAGAAGGT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTGTTTCCAAGAAAAGTTATAAGTTTTCTAACAAAATCTGTAAAACAGATAAAAGAAGGT 780

Qy    781 CGCCTCAAAGAGACACAAAAGCACCGAGTGGATTTCCTTCAGCTGATGATTGACTCTCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CGCCTCAAAGAGACACAAAAGCACCGAGTGGATTTCCTTCAGCTGATGATTGACTCTCAG 840

Qy    841 AATTCAAAAGACTCTGAGACCCACAAAGCTCTGTCTGATCTGGAGCTCATGGCCCAATCA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AATTCAAAAGACTCTGAGACCCACAAAGCTCTGTCTGATCTGGAGCTCATGGCCCAATCA 900

Qy    901 ATTATCTTTATTTTTGCTGGCTATGAAACCACGAGCAGTGTTCTCTCCTTCATTATATAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ATTATCTTTATTTTTGCTGGCTATGAAACCACGAGCAGTGTTCTCTCCTTCATTATATAT 960

Qy    961 GAACTGGCCACTCACCCTGATGTCCAGCAGAAAGTGCAGAAGGAAATTGATACAGTTTTA 1020
```

```
Db    961 GAACTGGCCACTCACCCTGATGTCCAGCAGAAAGTGCAGAAGGAAATTGATACAGTTTTA 1020

Qy   1021 CCCAATAAGGCACCACCCACCTATGATACTGTGCTACAGTTGGAGTATCTTGACATGGTG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 CCCAATAAGGCACCACCCACCTATGATACTGTGCTACAGTTGGAGTATCTTGACATGGTG 1080

Qy   1081 GTGAATGAAACACTCAGATTATTCCCAGTTGCTATGAGACTTGAGAGGGTCTGCAAAAAA 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GTGAATGAAACACTCAGATTATTCCCAGTTGCTATGAGACTTGAGAGGGTCTGCAAAAAA 1140

Qy   1141 GATGTTGAAATCAATGGGATGTTTATTCCCAAAGGGGTGGTGGTGATGATTCCAAGCTAT 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GATGTTGAAATCAATGGGATGTTTATTCCCAAAGGGGTGGTGGTGATGATTCCAAGCTAT 1200

Qy   1201 GTTCTTCATCATGACCCAAAGTACTGGACAGAGCCTGAGAAGTTCCTCCCTGAAAGGTTC 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTTCTTCATCATGACCCAAAGTACTGGACAGAGCCTGAGAAGTTCCTCCCTGAAAGGTTC 1260

Qy   1261 AGTAAAAAGAACAAGGACAACATAGATCCTTACATATACACACCCTTTGGAAGTGGACCC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGTAAAAAGAACAAGGACAACATAGATCCTTACATATACACACCCTTTGGAAGTGGACCC 1320

Qy   1321 AGAAACTGCATTGGCATGAGGTTTGCTCTCGTGAACATGAAACTTGCTCTAGTCAGAGTC 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGAAACTGCATTGGCATGAGGTTTGCTCTCGTGAACATGAAACTTGCTCTAGTCAGAGTC 1380

Qy   1381 CTTCAGAACTTCTCCTTCAAACCTTGTAAAGAAACACAGATCCCCCTGAAATTACGCTTT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CTTCAGAACTTCTCCTTCAAACCTTGTAAAGAAACACAGATCCCCCTGAAATTACGCTTT 1440

Qy   1441 GGAGGACTTCTTCTAACAGAAAAACCCATTGTTCTAAAGGCTGAGTCAAGGGATGAGACC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GGAGGACTTCTTCTAACAGAAAAACCCATTGTTCTAAAGGCTGAGTCAAGGGATGAGACC 1500

Qy   1501 GTAAGTGGAGCCTGATTTCCCTAAGGACTTCTGGTTTGCTCTTTAAGAAAGCTGTGCCCC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GTAAGTGGAGCCTGATTTCCCTAAGGACTTCTGGTTTGCTCTTTAAGAAAGCTGTGCCCC 1560

Qy   1561 AGAACACCAGAGACCTCAAATTACTTTACAAATAGAACCCTGAAATGAAGACGGGCTTCA 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AGAACACCAGAGACCTCAAATTACTTTACAAATAGAACCCTGAAATGAAGACGGGCTTCA 1620

Qy   1621 TCCAATGTGCTGCATAAATAATCAGGGATTCTGTACGTGCATTGTGCTCTCTCATGGTCT 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TCCAATGTGCTGCATAAATAATCAGGGATTCTGTACGTGCATTGTGCTCTCTCATGGTCT 1680

Qy   1681 GTATAGAGTGTTATACTTGGTAATATAGAGGAGATGACCAAATCAGTGCTGGGGAAGTAG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GTATAGAGTGTTATACTTGGTAATATAGAGGAGATGACCAAATCAGTGCTGGGGAAGTAG 1740

Qy   1741 ATTTGGCTTCTCTGCTTCTCATAGGACTATCTCCACCACCCCCAGTTAGCACCATTAACT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ATTTGGCTTCTCTGCTTCTCATAGGACTATCTCCACCACCCCCAGTTAGCACCATTAACT 1800

Qy   1801 CCTCCTGAGCTCTGATAACATAATTAACATTTCTCAATAATTTCAACCACAATCATTAAT 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CCTCCTGAGCTCTGATAACATAATTAACATTTCTCAATAATTTCAACCACAATCATTAAT 1860

Qy   1861 AAAAATAGGAATTATTTTGATGGCTCTAACAGTGACATTTATATCATGTGTTATATCTGT 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 AAAAATAGGAATTATTTTGATGGCTCTAACAGTGACATTTATATCATGTGTTATATCTGT 1920

Qy   1921 AGTATTCTATAGTAAGCTTTATATTAAGCAAATCAATAAAAACCTCTTTAC 1971
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 AGTATTCTATAGTAAGCTTTATATTAAGCAAATCAATAAAAACCTCTTTAC 1971
```

SEQ ID NO: 865

```
RESULT 2
AW967166/c
LOCUS       AW967166                659 bp    mRNA    linear   EST 01-JUN-2000
DEFINITION  EST379240 MAGE resequences, MAGJ Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW967166
VERSION     AW967166.1  GI:8157002
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 659)
  AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt
            ,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
            Quackenbush,J.
  TITLE     Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
  JOURNAL   Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 237
            Seq primer: Reverse.
FEATURES             Location/Qualifiers
     source          1. .659
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="MAGE resequences, MAGJ"
                     /note="Vector: pBluescriptSKm"
BASE COUNT      213 a    125 c    148 g    173 t
ORIGIN
```

Query Match 92.3%; Score 389.6; DB 10; Length 659;
Best Local Similarity 95.3%; Pred. No. 3.4e-72;
Matches 403; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

```
Qy      1 TGTTTTTTTTNGTATTTCTCATGTATACTTCATTTATTTTATTAATNANCNAANCCCTGT 60
          | ||||| || ||||||||||||||||||||||||||||||||||| | | || ||||||
Db    455 TTTTTTTCTTGGTATTTCTCATGTATACTTCATTTATTTTATTAATAAGCAAAGCCCTGT 396

Qy     61 AAGGGANTNCTTTGCCTAGTCNTCCGACTNTGNTTNATCTTCATCTTGACTAATCNGGAA 120
          |||||    |||||||||||| ||||||| || || ||||||||||||||||||| ||||
Db    395 AAGGGGAGCCTTTGCCTAGTCCTCCGACTCTGATTCATCTTCATCTTGACTAATCTGGAA 336

Qy    121 GTAACNAAG-TCGTAGGTCTCCTTGTCAGATGCAANCANTCGAAGCCAATCACGAAGATT 179
          ||||| ||| |||||||||||||||||||||||| || ||||||||||||||||||||
Db    335 GTAACGAAGTTCGTAGGTCTCCTTGTCAGATGCAACCACTCGAAGCCAATCACGAAGATT 276

Qy    180 GTTCTTCTTAAGGTATTTCTTGGTAAGGTATTTCAAATACCTTTTAGAGAACTGTTTCTC 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    275 GTTCTTCTTAAGGTATTTCTTGGTAAGGTATTTCAAATACCTTTTAGAGAACTGTTTCTC 216

Qy    240 AGAAACAACTGTGATTTTATTCTTGAAGCGTTCAATGTGAACAACATTCCCGAGATTTCC 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    215 AGAAACAACTGTGATTTTATTCTTGAAGCGTTCAATGTGAACAACATTCCCGAGATTTCC 156

Qy    300 AGTTTTGCCATTGACTTTAACCTTCTCCCGTAGAAATTGCTCAAAATTTCCAGAATCAAA 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    155 AGTTTTGCCATTGACTTTAACCTTCTCCCGTAGAAATTGCTCAAAATTTCCAGAATCAAA 96

Qy    360 AATTCCATCTTCTACTGGATGAGTAAGGTCCAAATTAAACCTCCAGGTTGACCTCTTGGG 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     95 AATTCCATCTTCTACTGGATGAGTAAGGTCCAAATTAAACCTCCAGGTTGACCTCTTGGG 36

Qy    420 CTT 422
          |||
Db     35 CTT 33
```

ALIGNMENTS

SEQ ID NO:1015

RESULT 1
N22479
LOCUS N22479 354 bp mRNA linear EST 20-DEC-1995
DEFINITION yw35a11.s1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:254204 3', mRNA sequence.
ACCESSION N22479
VERSION N22479.1 GI:1128613
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 354)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 268

Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 484 Std Error: 0.00
Seq primer: m13 -40 forward
High quality sequence stop: 268.

FEATURES Location/Qualifiers
source 1. .354
/organism="Homo sapiens"
/db_xref="GDB:3891657"
/db_xref="taxon:9606"
/clone="IMAGE:254204"
/clone_lib="Morton Fetal Cochlea"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned unidirectionally. Primer: Oligo dT. Fetal cochlea, normal. 37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP XR Vector. Library constructed by N. Robertson, C. Morton. -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"

BASE COUNT 120 a 68 c 37 g 129 t
ORIGIN

Query Match 100.0%; Score 354; DB 14; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.3e-46;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TTTTTTTTTTTTTTTTTGTCATATCCATAAAGTTTATTTTAAAAAACTGTTAAAAACTAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TTTTTTTTTTTTTTTTTGTCATATCCATAAAGTTTATTTTAAAAAACTGTTAAAAACTAA 60

Qy      61 CCATACAGGTTTATTAATATACTTAAAAAGTTTGTTCCCTATGTTGAAGTAAAATACATT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CCATACAGGTTTATTAATATACTTAAAAAGTTTGTTCCCTATGTTGAAGTAAAATACATT 120

Qy     121 AGCAACATCTTCCGGACACCATCTTTATAAAAGTAAAACTTCTAGATCCTGAAATGTACT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGCAACATCTTCCGGACACCATCTTTATAAAAGTAAAACTTCTAGATCCTGAAATGTACT 180

Qy     181 ACAGTAGAGTCTATAGTTTACACTTTTAATCACAGATTGGAATTCATTCTCCTTACTCCC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACAGTAGAGTCTATAGTTTACACTTTTAATCACAGATTGGAATTCATTCTCCTTACTCCC 240

Qy     241 CTACTTCCCACATGTGGCAGTTATTACTTCAAAATTAATGACATTCACTCATGTTATACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTACTTCCCACATGTGGCAGTTATTACTTCAAAATTAATGACATTCACTCATGTTATACT 300

Qy     301 ACCACAGATCCTTAAATAGAGTACATACTGCATAATTACTAACAGAGCCAGTCT 354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ACCACAGATCCTTAAATAGAGTACATACTGCATAATTACTAACAGAGCCAGTCT 354
```

SEQ ID NO: 1027

RESULT 4
AAZ51562/c
ID  AAZ51562 standard; cDNA; 580 BP.
XX
AC  AAZ51562;
XX
DT  21-JUN-2000  (first entry)
XX
DE  Human hypoxia response regulating gene, 77H4 related cDNA clone 18E.
XX
KW  Hypoxia response regulating gene; gene 77H4; human; EST 18E; cardiant;
KW  apoptosis; angiogenesis; vasotropic; cytostatic; ophthalmological;
KW  cerebroprotective; antagonist; regulator; inhibitor; treatment; tumour;
KW  hypoxia associated pathology; HAP; gene therapy; diagnosis; ischaemia;
KW  steroid receptor coactivator; SRA; retinopathy; myocardial infarction;
KW  stroke; ss.
XX
OS  Homo sapiens.
XX
FH  Key             Location/Qualifiers
FT  polyA_signal    536..541
FT                  /*tag= a
XX
PN  WO200012525-A1.
XX
PD  09-MAR-2000.
XX
PF  27-AUG-1999;  99WO-US20394.
XX
PR  27-AUG-1998;  98US-0098158.
PR  05-MAY-1999;  99US-0132684.
XX
PA  (QUAR-) QUARK BIOTECH INC.
PA  (KOHN/) KOHN K.
XX
PI  Einat P,  Skaliter R,  Feinstein E;
XX
DR  WPI; 2000-256577/22.
XX
PT  Novel polynucleotides capable of regulating angiogenesis or apoptosis
PT  useful for diagnosis and treatment of hypoxia, ischemia and tumor
PT  growth -
XX
PS  Claim 1; Fig 7a; 78pp; English.
XX
CC  The present sequence is the human hypoxia response regulating gene,
CC  77H4, related cDNA clone 18E. The gene 77H4 has similarity to steroid
CC  receptor transcriptional co-activator, SRA function and can serve as
CC  a coactivator in some transcriptional complexes. It has vasotropic,
CC  cardiant, ophthalmological, cytostatic and cerebroprotective activity.
CC  Antagonist of the encoded protein, functions as a regulator of apoptosis
CC  or angiogenesis. The protein encoded by this polynucleotide, the
CC  biologically active product from enzymatic activity of the protein or
CC  inhibitor of the enzymatic activity is useful for regulating hypoxia
CC  associated pathologies (HAP). It is useful for gene therapy, diagnosis
CC  and treatment of tumour growth and ischaemia, e.g., retinopathy,
CC  myocardial infarction and stroke.
XX
SQ  Sequence 580 BP; 177 A; 111 C; 134 G; 158 T; 0 other;

Query Match  97.5%;  Score 428;  DB 21;  Length 580;
Best Local Similarity  99.8%;  Pred. No. 1.3e-125;
Matches 439;  Conservative  0;  Mismatches  0;  Indels  1;  Gaps  1;

```
Qy      1 TTTTTGATTTAAAAGATTTTATTTTCTTTATGCAGGTAGGCAGTTAGAAATTTCAAAGTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    558 TTTTTGATTTAAAAGATTTTATTTTCTTTATGCAGGTAGGCAGTTAGAAATTTCAAAGTC 499

Qy     61 TAACAATGACATTCTTGAAGTGGGCACAG-CTTTTAAACTCAGGCTATGTATACAGTAAC 119
          |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    498 TAACAATGACATTCTTGAAGTGGGCACAGCCTTTTAAACTCAGGCTATGTATACAGTAAC 439

Qy    120 CTTGTGGAACTGGTTCAGCCAGATCTTCACTTTCATGAAAGCACAGGGTCTGTCCTTTTC 179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    438 CTTGTGGAACTGGTTCAGCCAGATCTTCACTTTCATGAAAGCACAGGGTCTGTCCTTTTC 379

Qy    180 TTTCCAGAGGGCTCCTCTCATATTCCATCGCCAGTTTCTGTTACAAGGCAGACTGAATCA 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    378 TTTCCAGAGGGCTCCTCTCATATTCCATCGCCAGTTTCTGTTACAAGGCAGACTGAATCA 319

Qy    240 AGCCAAGATCAACACACACTGGTACACGTGGCTCCCAACCAATTTTATATGTATATATAT 299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 AGCCAAGATCAACACACACTGGTACACGTGGCTCCCAACCAATTTTATATGTATATATAT 259

Qy    300 ATTCTACTTCCAACACCCGCATTCATCCTGGTTCAATCAAAGCCTGGTTTTGGCCAACAA 359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 ATTCTACTTCCAACACCCGCATTCATCCTGGTTCAATCAAAGCCTGGTTTTGGCCAACAA 199

Qy    360 TAAACTCGTCAGGAGATCGAAGGTTGTAGATGTCTGCACGTGGCTTCCTTGGAGGTCCAG 419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 TAAACTCGTCAGGAGATCGAAGGTTGTAGATGTCTGCACGTGGCTTCCTTGGAGGTCCAG 139

Qy    420 TGGTGACTCCCTCTTCCAAA 439
          ||||||||||||||||||||
Db    138 TGGTGACTCCCTCTTCCAAA 119
```

SEQ ID NO: 1027

ALIGNMENTS

```
RESULT 1
N98464
LOCUS       N98464                   439 bp    mRNA    linear   EST 20-AUG-1996
DEFINITION  zb86b06.s1 Soares_senescent_fibroblasts_NbHSF Homo sapiens cDNA
            clone IMAGE:310451 3', mRNA sequence.
ACCESSION   N98464
VERSION     N98464.1  GI:1269909
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 439)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 568    Std Error: 0.00
            Seq primer: mob.REGA+ET
            High quality sequence stop: 377.
FEATURES             Location/Qualifiers
     source          1. .439
                     /organism="Homo sapiens"
                     /db_xref="GDB:1253865"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:310451"
                     /clone_lib="Soares_senescent_fibroblasts_NbHSF"
                     /tissue_type="senescent fibroblast"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Vector: pT7T3D (Pharmacia) with a modified
                     polylinker V_TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
                     ; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      118 a    101 c     83 g    137 t
ORIGIN

  Query Match          100.0%;  Score 439;  DB 14;  Length 439;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-110;
  Matches  439;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TTTTTGATTTAAAAGATTTTATTTTCTTTATGCAGGTAGGCAGTTAGAAATTTCAAAGTC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TTTTTGATTTAAAAGATTTTATTTTCTTTATGCAGGTAGGCAGTTAGAAATTTCAAAGTC 60

Qy      61 TAACAATGACATTCTTGAAGTGGGCACAGCTTTTAAACTCAGGCTATGTATACAGTAACC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TAACAATGACATTCTTGAAGTGGGCACAGCTTTTAAACTCAGGCTATGTATACAGTAACC 120

Qy     121 TTGTGGAACTGGTTCAGCCAGATCTTCACTTTCATGAAAGCACAGGGTCTGTCCTTTTCT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTGTGGAACTGGTTCAGCCAGATCTTCACTTTCATGAAAGCACAGGGTCTGTCCTTTTCT 180

Qy     181 TTCCAGAGGGCTCCTCTCATATTCCATCGCCAGTTTCTGTTACAAGGCAGACTGAATCAA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTCCAGAGGGCTCCTCTCATATTCCATCGCCAGTTTCTGTTACAAGGCAGACTGAATCAA 240

Qy     241 GCCAAGATCAACACACACTGGTACACGTGGCTCCCAACCAATTTTATATGTATATATATA 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GCCAAGATCAACACACACTGGTACACGTGGCTCCCAACCAATTTTATATGTATATATATA 300

Qy     301 TTCTACTTCCAACACCCGCATTCATCCTGGTTCAATCAAAGCCTGGTTTTGGCCAACAAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTCTACTTCCAACACCCGCATTCATCCTGGTTCAATCAAAGCCTGGTTTTGGCCAACAAT 360

Qy     361 AAACTCGTCAGGAGATCGAAGGTTGTAGATGTCTGCACGTGGCTTCCTTGGAGGTCCAGT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAACTCGTCAGGAGATCGAAGGTTGTAGATGTCTGCACGTGGCTTCCTTGGAGGTCCAGT 420

Qy     421 GGTGACTCCCTCTTCCAAA 439
          ||||||||||||||||||||
Db     421 GGTGACTCCCTCTTCCAAA 439
```

SEQ ID NO: 1051

ALIGNMENTS

RESULT 1
W90146
LOCUS W90146 304 bp mRNA linear EST 07-MAY-1997
DEFINITION zh75e03.s1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:417916 3', mRNA sequence.
ACCESSION W90146
VERSION W90146.1 GI:1406136
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 304)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1214 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 267.

FEATURES Location/Qualifiers
source 1. .304
/organism="Homo sapiens"
/db_xref="GDB:1326386"
/db_xref="taxon:9606"
/clone="IMAGE:417916"
/clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen 1NFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer.[5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 86 a 72 c 45 g 100 t 1 others
ORIGIN

Query Match 99.7%; Score 303; DB 14; Length 304;
Best Local Similarity 100.0%; Pred. No. 5e-75;
Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CACTTATAGCCAATATTTAATAATCCCATATTAACTGATGTGTAAAAAATGTCTTTATGA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CACTTATAGCCAATATTTAATAATCCCATATTAACTGATGTGTAAAAAATGTCTTTATGA 60

Qy     61 TCTGTTACCACCCAAAAGAATGCATCATAACTTTCAAGANTATGTTCTTTGACTTCTAAC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTGTTACCACCCAAAAGAATGCATCATAACTTTCAAGANTATGTTCTTTGACTTCTAAC 120

Qy    121 CTCTGCTCTTCTTTAGAATTACCTTTGCTGCGGCCAGTACATGCTCCTTGTTAATGACTC 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTCTGCTCTTCTTTAGAATTACCTTTGCTGCGGCCAGTACATGCTCCTTGTTAATGACTC 180

Qy    181 TACATTTACTCGCACAAGCGTTTGTCCTGGACTCTTCTGCTAATCGATGAACAAACAGTA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TACATTTACTCGCACAAGCGTTTGTCCTGGACTCTTCTGCTAATCGATGAACAAACAGTA 240

Qy    241 AACAGTTCAGATGGACCAATAAGTCACCACTTTTCTCCAGACCGAAGTTGGAGGGCTTCT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AACAGTTCAGATGGACCAATAAGTCACCACTTTTCTCCAGACCGAAGTTGGAGGGCTTCT 300

Qy    301 TTCC 304
          ||||
Db    301 TTCC 304
```